Journal of Experimental Social Psychology 50 (2014) 175–183

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Easy on the eyes, or hard to categorize: Classification difficulty decreases the appeal of facial blends

Jamin Halberstadt [a,*], Piotr Winkielman [b,*]

[a] Department of Psychology, University of Otago, Dunedin, New Zealand
[b] Department of Psychology, University of California, San Diego, CA, USA

HIGHLIGHTS

- Facial attractiveness is partially due to the ease with which faces can be categorized.
- The attractiveness of face morphs is reduced when participants first classify the faces.
- Bi-racial faces are less attractive when they are first classified by race.
- Participants smile less at cross-race faces after classifying them by race.

ARTICLE INFO

Article history:
Received 9 April 2013
Revised 12 August 2013
Available online 25 August 2013

Keywords:
Social cognition
Facial attractiveness
Processing fluency
Cross-race faces
Social categorization
EMG

ABSTRACT

Social information processing often involves categorization. When such categorization is difficult, the disfluency may elicit negative affect that could generalize to a variety of stimulus judgments. In the current studies we experimentally apply this theoretical analysis to two classic and highly socially relevant facial attractiveness phenomena: the beauty-in-averageness effect and the appeal of bi-racial faces. Studies 1 and 2 show that same-race (Caucasian–Caucasian) morphs are rated as more attractive than the individual faces composing them — a classic "beauty-in-averageness effect." Critically, however, this effect is reduced or eliminated when participants first classify the faces in terms of their "parents," and only if that classification is difficult. Studies 3 and 4 extend these results to show that classifying bi-racial individuals in terms of their racial identity reduces perceivers' ratings of attractiveness and reverses perceivers' tendency to smile at them, as measured by facial electromyography (EMG). Together, these four studies support the proposal that facial attractiveness is partially a function of the experience of social categorization, and that such experience depends critically on the nature of the categories into which an individual can be classified.

Introduction

Social psychologists have long been interested in attractiveness, and for good reason. Understanding what makes people and things "attractive" informs about the basic operation of our affective system and its interactions with cognition (Berntson & Cacioppo, 2009; Zajonc, 1998). Of course, social psychologists also care about attractiveness because it plays a significant role in our society, influencing — and being influenced by — a variety of social inferences (Eagly, Ashmore, Makhijani, & Longo, 1991; Etcoff, 2000). In that context, social psychologists are particularly interested in cognitive processes that can change seemingly "objective" evaluations. An important and well-studied example is social categorization. For example, some of the appeal of category exemplars comes from the group to which a person belongs, such that the same person is more appealing when seen as a member of a positive versus a negative group (Fiske, 1982). In this case, category information provides top-down input into information search and integration that biases overall evaluation.

In the current set of studies, however, we explore a very different mechanism by which categorization influences attractiveness—the *ease* with which a person can be categorized, independent of category valence. We show that the same face may be more or less attractive depending on how *difficult* it is to classify, which in turn depends on the salience of competing category memberships. We first illustrate these effects in a classic attractiveness phenomenon, the attractiveness of morphed faces (the "beauty-in-averageness effect"), and then extend our analysis to "real" morphs: bi-racial faces.

The beauty-in-averageness effect

Among the many fascinating phenomena associated with faces, one of the most robust and nonobvious is the aesthetic impact of blending

* Corresponding authors.
E-mail addresses: jhalbers@psy.otago.ac.nz (J. Halberstadt), pwinkielman@ucsd.edu (P. Winkielman).

0022-1031/$ – see front matter
http://dx.doi.org/10.1016/j.jesp.2013.08.004

them, the so-called "beauty in averageness" effect. First observed by Galton (1879), and reintroduced to the attractiveness literature by Langlois and Roggman (1990), the effect refers to the fact that "average" faces, produced by blending a number of unmodified component images, are judged as more attractive than the faces used to create them (Halberstadt, 2006; Langlois & Roggman, 1990; Rhodes & Tremewan, 1996; Rubenstein, Kalakanis, & Langlois, 1999). Generally, adding more faces to a composite makes them increasingly more attractive, and faces can be made more (or less) attractive by distorting them toward (or away from) a population average (Rhodes & Tremewan, 1996).

There probably is not one single mechanism for the beauty-in-averageness effect (see Halberstadt, 2006, for some discussion). However, one contributing variable may be *processing fluency* — the speed and ease of perceptual and conceptual mental operations associated with the stimulus (Winkielman, Schwarz, Fazendeiro, & Reber, 2003). Processing fluency is often associated with positive affect — either because the experience of successful identification or classification is pleasant in itself, or because it signals something positive about the stimulus, such as its safety or familiarity (see Winkielman et al., 2003, for review). Perceivers report increased liking for stimuli, such as drawings of everyday objects, when the fluency of stimulus identification or stimulus categorization has been experimentally enhanced (e.g., by priming or perceptual clarification, Reber, Winkielman, & Schwarz, 1998). Converging physiological evidence of these positive reactions comes from studies using facial electromyography (EMG), which found that incipient activity of smiling-related muscles increases in response to fluent stimuli (e.g., Winkielman & Cacioppo, 2001; Winkielman, Halberstadt, Fazendeiro, & Catty, 2006).

Though it has never been directly shown, part of the appeal of human face composites may be due to the ease with which they can be identified and classified. This suggestion is consistent with research using non-social categories, which reported a link between the ease of classifying an image and its attractiveness (Winkielman et al., 2006). In that research, participants saw dot patterns that varied in similarity to a randomly generated prototype. Participants' task was either to quickly classify the patterns into their respective categories or to rate their attractiveness. Results showed that the patterns most similar to the prototype were easier to classify and most preferred. Conversely, distorted patterns (i.e., the ones least similar to the original prototype) were most difficult to classify and least preferred, as reflected in both self-report of attractiveness and physiological measures of affect. Furthermore, the effect of distortion on attractiveness was partially mediated by their classification (dis)fluency.

However, there is an apparent paradox inherent in the fluency account of the preference for face composites (which we also refer to as morphs, averages, or blends). On the one hand, composite stimuli (averages) should be easy to process, because they represent a good summary of the perceiver's previous experience (i.e., a category prototype). On the other hand, composite stimuli should be difficult to process, because they are maximally ambiguous with regard to the original faces composing them. The key to resolving this paradox, we propose, lies in an appreciation of the categorical relativity of fluency effects. Note that composite faces are both facial prototypes *and* facial distortions, depending on the current category of which they are judged a member (i.e., whether they are treated as examples of the general category of "faces," or as examples of the specific component faces). Moreover, to the extent that processing fluency influences facial attractiveness, the same facial composites should be relatively attractive when they are easy to process (because they represent the central tendency of the faces to which a perceiver has been exposed), and less attractive when they are difficult to process (because they are highly ambiguous as to their identity). Consistent with this proposal, we recently found that morphs of two celebrities are more attractive than their original component faces (the "beauty-in-averageness" effect) *only* when the celebrities are unknown in the country where participants were tested (i.e., they are famous in a different country). When celebrities are well-known in the participants' country, blends of them are *less* attractive than the originals (an "unattractiveness-in-averages" effect). This reversal can be theoretically explained by assuming that the salience of competing identities makes the morphed celebrities difficult to process and classify (Halberstadt, Pecher, Zeelenberg, Ip Wai, & Winkielman, 2013). Importantly though, this study did not experimentally manipulate categorization; it relied on participants' extensive real world experience (or lack thereof) with individual "parent" faces. Further, this study did not assess or manipulate fluency. As such, it can only be interpreted as consistent with the proposed theoretical account, in which fluency serves as a mechanism underlying attractiveness judgments.

Note that the same insight regarding the categorical relativity of fluency and attractiveness can be applied to resolve the puzzle of "real" morphs — mixed-race faces. This is important because the population of bi- and multi-racial individuals is growing. According to 2010 Census Bureau estimates, since 2000 (the first year Americans were allowed to check one or more races on the survey) the overall population of individuals identifying with mixed-race has grown by roughly 35 percent, and the population of children by roughly 50% (Saulny, 2011a, 2011b). Yet, our understanding of how such bi- and multi-racial individuals are perceived and evaluated lags behind (Campbell & Herman, 2010). For example, of the few studies examining affective judgments of mixed-race individuals, some report more positive reactions and some more negative reactions. More specifically, Lewis (2010) found that mixed-race individuals are perceived as more attractive than uni-racial individuals, and presumably benefit from "halo effects," such as attributions of warmth and competence (Eagly et al., 1991). However, Sanchez and Bonam (2009) found that mixed-race job applicants were actually judged as *less* warm and competent than uni-racial applicants, and many qualitative reports conclude that bi-racial individuals are especially vulnerable to racism, social isolation, and romantic rejection (e.g., Brandell, 1988; Gibbs & Moskowitz-Sweet, 1991; Jackman, Wagner, & Johnson, 2001; Winn & Priest, 1993). From the current perspective, such findings are not necessarily contradictory, because the appeal of mixed-race faces should depend on how — and how easily — they can be classified. Like laboratory morphs, these faces are both prototypical of the larger set of faces to which the perceiver has been exposed, and atypical examples of the individual constituent racial groups. Therefore, their fluency, and in turn their attractiveness, should depend on whether they are judged with reference to their race. A man with both Chinese and Caucasian features, for example, should be a relatively attractive "man," but a relatively unattractive "Asian Man" or "Caucasian Man."

Current studies

The goal of the current set of studies was to test the sensitivity of the "beauty-in-averageness" effect to changes in category structure, which would thereby experimentally implicate processing fluency in the attractiveness of facial blends. In Studies 1 and 2, participants rated the attractiveness of uni-racial (Caucasian–Caucasian) morphs under conditions that required implicit classification in terms of their parent faces (or under control conditions that did not). Studies 3 and 4 conceptually replicate the paradigm using bi-racial (Caucasian–Asian) morphs and include psychophysiological measures of affect. For both stimulus sets, we expected to replicate the "beauty-in-averageness" effect in the control conditions, such that morphs would be rated as more attractive and would elicit more positive emotional expressions than their parents (whether of the same or different races). However, when attractiveness judgments are preceded by classification in terms of the parents, the advantage for morphs should be weakened (or even reversed) due to the disfluency associated with the ambiguity of morphed images.

Study 1

Method

Participants and stimuli

Seventy-seven female and 28 male University of Otago students participated as part of a course research requirement or for extra credit. The stimuli were two sets (for purposes of replication) of 26 digitally blended faces of young Caucasian women. Each set was created by mapping corresponding points of the two original images and then morphing one into the other in 4% intervals using Morph 2.5 software (for details of this well-established technique of stimulus development, see, for example, Rhodes, Sumich, & Byatt, 1999). Thus, the blends represented mathematically equal steps between the two original images. Two "families" were identified in each set, representing the 13 blends to each side of the morphed midpoint (the 50–50 blend of the two original faces). Examples appear in Fig. 1.

Procedure

Participants gave informed consent and then were given all instructions and stimuli on 17-inch iMac computers located in sound- and light-attenuated experimental cubicles. Participants were randomly assigned to judge one of the two sets of stimulus faces. All participants received the following explanation of the experiment:

> Learning about social groups is critical to effective social interaction. For example, suppose you're going to meet a new flatmate. If you know the person's family ahead of time, you can use this information to help evaluate the person and judge their behaviour. In this experiment, we are studying how people learn and use social categories.

Participants were then asked to study for 30 s "members of two different families" (actually the parent faces used to create the stimulus blends), presented simultaneously on the left and right of the computer screen (randomly determined), with the "family names" "Ack" and "Blub" respectively presented below them. Afterwards, participants were informed that they would next "see some female faces" and be asked to decide whether each belongs to the Ack or Blub family. Participants were instructed to respond as quickly as possible, without making errors, by pressing the "Z" or "/" keys, appropriately labelled. Feedback was given on all trials, in the form of a brief "beep" and "buzz" auditory cue (correct and incorrect family classifications, respectively). The 26 stimulus faces each appeared individually in the center of the screen, three times over in consecutive randomised blocks. The stimulus face disappeared as soon as the participant made a response,

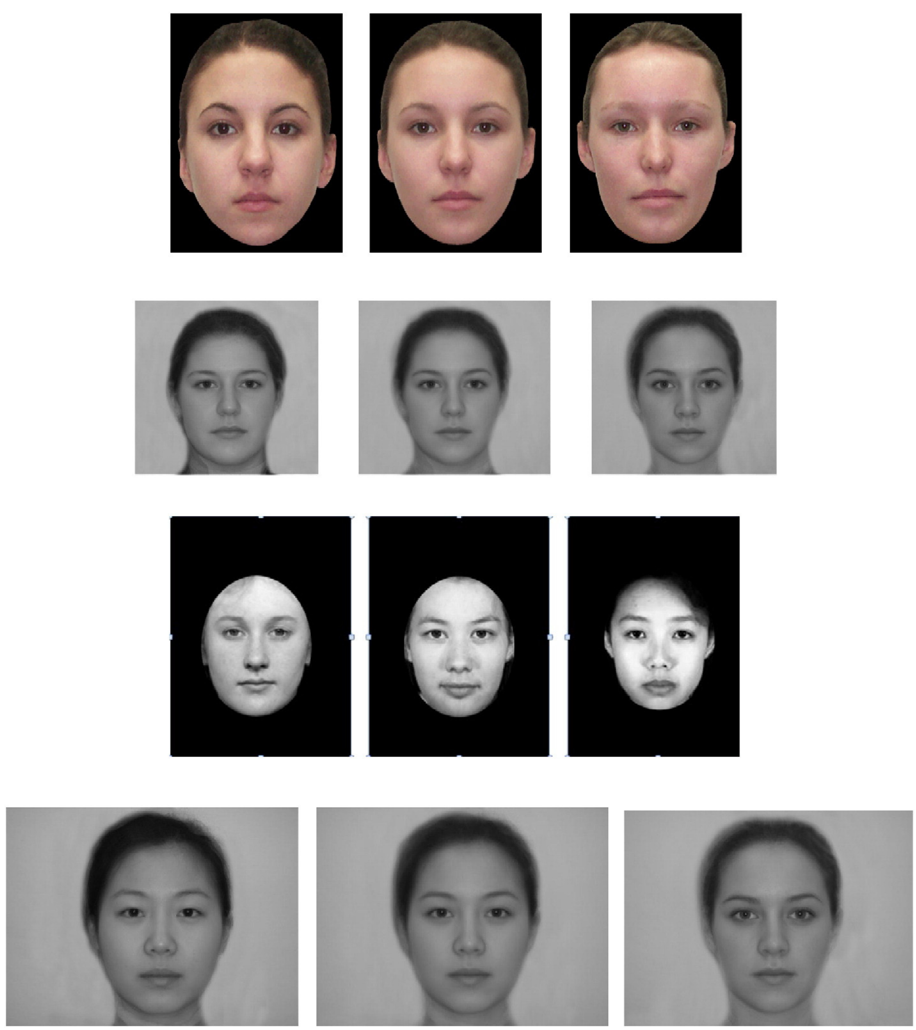

Fig. 1. Examples of stimuli used in Studies 1, 2, 3, and 4 (top to bottom). In rows 1, 2, and 4, the middle face is a 50% blend of the parent faces on either side of it.

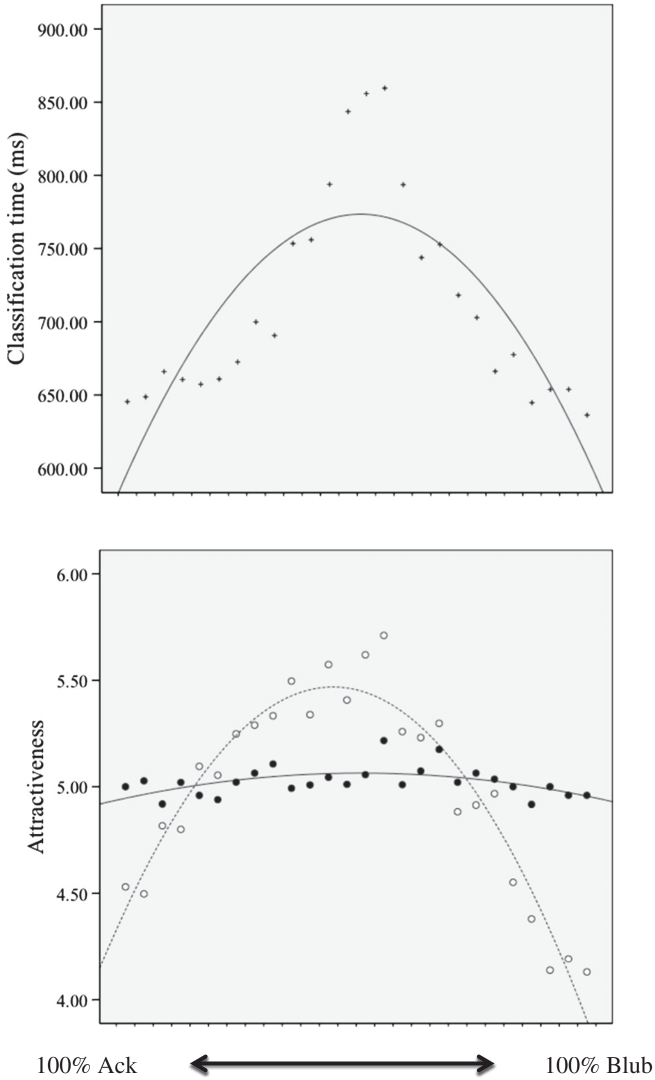

Fig. 2. Mean classification time (top panel) and attractiveness of 26 blends in Experiment 1, as a function of morph level and experimental condition. The two-scale condition, which requires implicit classification of the faces prior to rating them, is represented by the solid line.

which was then followed by a 1000 ms interval before the presentation of the next face.

After classifying all stimulus faces, participants repeated the study session before being informed of our additional interest in "what makes people seem attractive or unattractive" and asked to rate the attractiveness of "members of the Ack and Blub families." In the control condition, participants were presented with a single sliding scale anchored at 1 ("very unattractive") to 9 ("very attractive"). In the experimental condition, participants were presented with two scales labelled "Ack Scale" and "Blub Scale," with additional instructions to make their attractiveness rating on the appropriate scale. Note that this task implicitly requires categorization of the face (via scale selection), before the attractiveness judgment is made. In both scale conditions, all 26 faces appeared individually in the center of the screen, in random order, and each face was replaced by the next as soon as a participant responded.

Results and discussion

Preprocessing

Following Winkielman et al. (2006), trials on which participants assigned faces to the wrong family, or on which participants responded extremely quickly or slowly (less than 200 ms or more than 2183 ms, three *SDs* above the mean of 744 ms) were not analyzed. The remaining 97% of the data were collapsed across participants to create average fluency and attractiveness estimates at each of 26 morph levels in each experimental condition, which are plotted in Fig. 2.[1] The primary analyses in Studies 1 and 2 were conducted at the level of the stimulus.

Main analyses

As seen in Fig. 2, morphing the parent faces made them significantly disfluent, as evidenced by a strong and symmetric quadratic relationship, between morph level and classification time, with the most strongly blended morphs (levels 13 and 14) classified more than 200 ms slower than the parent faces (levels 1 and 26). The fluency effects were tested formally with a multiple regression, in which morph level and squared morph level (representing the linear and quadratic effects), and their respective interactions with scale condition, were used to predict classification fluency. The only significant predictor was the quadratic main effect of morph level, Beta $= -.78$, $F(1,46) = 74.75$, $p < .001$. As expected (because participants performed the classification task before the scale manipulation was introduced), the morphing–disfluency relationship did not vary by experimental condition ($p > .5$).

The same regression analysis on attractiveness ratings revealed significant effects of all predictors but, most pertinent to our hypotheses, a strong interaction between scale condition and squared morph level (the quadratic component), Beta $= .83$, $F(1,46) = 144.95$, $p < .001$. The interaction was due to the fact that, as predicted, a quadratic model fits the attractiveness data far better in the one scale condition (r-squared $= .92$) than in the two scale condition (r-squared $= .29$). As seen in Fig. 2, ratings in the control condition replicated the classic beauty-in-averageness effect: morphs were increasingly attractive as a function of the degree to which they were blended, and the most extreme blends were more than a full rating point more attractive than the original faces that created them. However, implicitly requiring participants to classify the stimuli (in order to use the appropriate rating scale) effectively "flattened" the morphing–attractiveness relationship, such that less blended faces were relatively more attractive, and more blended faces relatively less attractive, than in the control condition.

In order to test the hypothesized role of classification (dis)fluency in the flattening effect, we correlated the mean classification speed of each stimulus (measured prior to the experimental manipulation) with the degree to which classification reduces its attractiveness (i.e., the difference between its attractiveness in the classification versus the control conditions). This correlation was highly significant, $r(26) = .75$, $p < .001$. In other words, when attractiveness ratings involved implicit classification (the two scale condition), increasingly blended faces became increasingly less attractive (compared to control ratings) as a linear function of the speed with which their family membership could be identified.

Study 2

One possible alternative explanation of Study 1 is that the changes in the faces' attractiveness are not due to the effort required to classify them, but are instead an artifact of rating them on two scales rather than one. For example, participants might have found the use of two scales more confusing, leading to greater error or lower confidence in their ratings. Alternatively, two scales may have encouraged the use of

[1] Preliminary analyses indicated no effect of stimulus set on classification times, attractiveness ratings, or the correlation between them, and analyses were therefore collapsed across this variable.

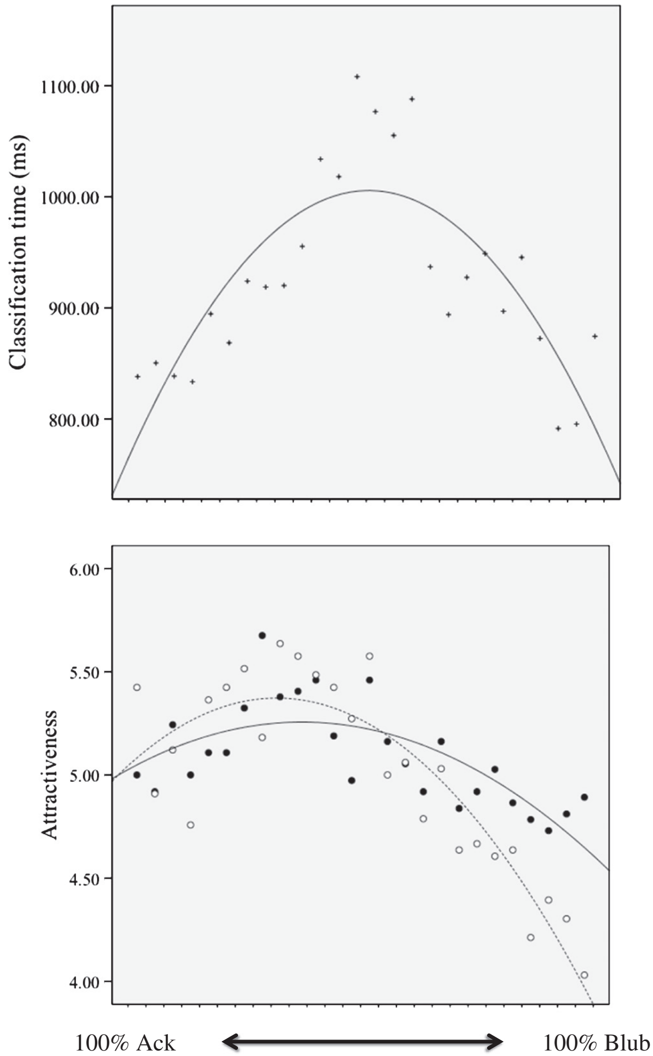

Fig. 3. Mean classification time (top panel) and attractiveness of 26 blends in Experiment 2, as a function of morph level and experimental condition. The effortful classification condition is represented by the solid line.

different, within-family standards of attractiveness, rendering differences between conditions difficult to interpret.

To verify that changes in the attractiveness of blends were indeed driven by the difficulty of classifying them, we conducted a replication which required *all* participants to judge face morphs on two family-specific scales. However, we varied the cognitive difficulty involved in doing so by informing half of the participants, on each trial, of the family to which a target belonged. Thus, both rating conditions were equated in terms of their use of family-specific categories and two rating scales (and in turn the standards for those scales). The conditions varied only in whether the participants classified the faces themselves (which should be difficult for blended faces) or whether this was done for them (removing the experience of difficulty).

Method

Participants and stimuli

Participants were 70 University of Otago student volunteers recruited through a job clearinghouse on campus, and reimbursed approximately US$10 to cover their travel expenses.

Stimuli were again 26 morphs, derived from two separate 2-parent sets. However, in order to provide an even more stringent test of our hypotheses, the parents used in this case were themselves two-face (Caucasian–Caucasian) morphs (see Fig. 1). This reduces the influence of morphing artifacts (e.g., facial smoothing), which may artificially enhance the strength of the beauty-in-averageness effect.

Procedure

The procedure was identical to that described in Study 1, with the following exceptions. First, for greater power, morph set was a within subjects factor. Specifically, participants completed two replications of the study phase, and two replications of the rating phase, once for each stimulus set.[2] Second, after studying the parents, participants rated their blends under one of two conditions. One was the original two-scale condition described above, which requires implicit classification by the participants. In the new condition, there were also two scales, but participants were informed of the family to which a stimulus belonged (e.g., “This is a Blub”), and therefore which of the two scales should be used to rate it.

Results and discussion

Mean fluency and attractiveness ratings at each morph level were computed as in Study 1, and they appear in Fig. 3. As in that study, morphing the parents made them significantly disfluent, with the most strongly blended morphs classified more than 300 ms slower than the parent faces. A multiple regression confirmed, again, that the only significant predictor of classification fluency was the quadratic main effect of morph level, Beta $= -.78$, $F(1,46) = 71.81$, $p < .001$. As expected, experimental condition did not have any effect, since the classification task was performed before the scale manipulation was introduced.

The same analysis on attractiveness ratings replicated the critical interaction between experimental condition and quadratic (squared) morph level, Beta $= -.24$, $F(1,46) = 4.56$, $p < .05$. The interaction was due to the fact that, as predicted, a quadratic model fit the attractiveness data better when family membership was provided by the experimenter (r-squared $= .81$) than when it was generated by the participant (r-squared $= .52$). Furthermore, the interaction can again be statistically explained by classification fluency, with the difference between the two conditions' attractiveness ratings at each morph level correlated with the speed with which the morphed images could be classified in part 1 of the study, $r(26) = .38$, $p = .082$.

Study 2 replicated the basic findings of Study 1 and, critically, showed that the moderation of the beauty-in-averageness effect is not just a function of potentially different standards, or general confusion associated with using two scales. In Study 2, when the experimenter revealed a target's family membership, participants judged faces as more attractive as a function of the degree to which they were morphed, even when they made those judgments on family-specific scales. However, when they were required to judge family membership for themselves, the advantage for blends was relatively diminished, albeit not as dramatically as in Study 1. We propose that part of the decline in the beauty-in-averageness effect is due to disfluency. That is, the cognitive effort required to classify intermediate morphs, prior to rating them, produces negative affect that generalizes to the stimuli themselves.

Study 3

Studies 1 and 2 illustrate not only the causal importance of processing fluency in facial attractiveness, but also the context-specificity of both variables. We believe the same model can provide insight into

[2] Preliminary analyses revealed that, unexpectedly, only one of the two pairs of parent faces showed a pronounced beauty-in-averageness effect (the other showed a primarily linear relation between morphing and attractiveness, due apparently to the unexpectedly strong attractiveness of one parent), and the main analyses are confined to the blends of that pair.

the highly topical issue of how "real" facial morphs — i.e., mixed-race faces — are perceived. As noted in the Introduction, the limited research on judgments of individuals from bi- or multi-racial backgrounds is contradictory, with some findings suggesting more positive evaluations, and some more negative evaluations. In principle, the current theoretical framework can integrate both types of findings: Applying the logic of context-specific fluency, we hypothesize that the appeal of bi-racial individuals depends on whether or not they are processed in terms of the racial categories that compose them. Recall that in Studies 1 and 2, intermediate Caucasian–Caucasian morphs were relatively dissimilar to their parents, and therefore relatively disfluent and unattractive when classification in terms of the parents was required. By the same logic, bi-racial individuals are relatively dissimilar to prototypes of their constituent races, and so should be relatively unattractive when they are classified with respect to those groups. However, when race is *not* considered, mixed-race faces should be *more* attractive than the single-race faces (or at least not less attractive), given that they are relatively similar to the prototype of all encountered face exemplars (cf. Potter & Corneille, 2008).

Method

Participants and stimuli

Sixty-two (30 male) Caucasian undergraduates at the University of Otago participated to fulfill a research component of their psychology class. Stimuli consisted of two-face morphs created from 24 photographs of Chinese individuals and 24 photographs of Caucasian individuals, taken from the University of Western Australia's "Facelab" database (Rhodes, personal communication). Using morphing software (Rubley, 2010), each stimulus was blended with one same-race image and one other-race image to create 12 Chinese–Chinese morphs, 12 Caucasian–Caucasian morphs, and 24 Chinese–Caucasian morphs. All images were mounted on 360 × 480 black backgrounds with 225 × 275 oval masks. Examples appear in Fig. 1.

Procedure

Participants rated all 48 stimulus faces, presented individually in random order, on a 1 ("not attractive") to 9 ("very attractive") scale. Prior to rating each face, half of the participants categorized "as quickly and accurately as possible" whether an individual was "Caucasian" or "Asian," by pressing the "S" and "L" keyboard keys, respectively. The other half of participants categorized (using the same keys) each individual on a race-irrelevant dimension — whether each individual was "feeling positive" or "feeling negative" at the time that they were photographed (the individuals had ostensibly provided emotion self-reports). Each face appeared in the center of the screen for 3000 ms, following a 500 ms fixation cross (+). As soon as the participant responded, the face was replaced by the attractiveness scale. The ITI was 5 s.

Results and discussion

One classification time was faster than 200 ms, and thus was excluded from all analyses. Paired sample *t*-tests showed that classification times did not differ between Chinese–Chinese and Caucasian–Caucasian morphs in either experimental condition, so these data were combined into an average for single-race blends. A 2 (morph type: single-race versus mixed-race) × 2 (racial versus control classification) mixed-model ANOVA revealed main effects of both factors. Participants overall were slower to classify faces by emotion than by race (1629 ms versus 1375 ms, SEs = 51 ms and 50 ms), $F(1,60) = 12.64$, $p = .001$, $\eta_p^2 = .17$. Participants were also slower to classify mixed-race than within-race faces (1574 ms versus 1430 ms, SEs = 39 ms and 36 ms), $F(1,60) = 37.47$, $p < .001$, $\eta_p^2 = .38$. More importantly, a significant interaction, illustrated in Fig. 4, confirmed that mixed-race morphs were disfluent *only* in the context of the race classification task, $F(1,60) = 90.95$, $p < .001$, $\eta_p^2 = .60$. Specifically, participants were *slower* to judge the race of mixed-race individuals than of within-race individuals $t(31) = 10.46$, $p < .001$, $\eta_p^2 = .78$, but *faster* to judge the emotion of mixed-race than within-race individuals, $t(29) = -2.61$, $p < .05$, $\eta_p^2 = .19$.

A planned contrast tested the primary hypothesis that racial classification would decrease the beauty of mixed-race faces relative to within-race faces. This test was significant, $t(60) = 1.74$, $p < .05$, $\eta_p^2 = .05$ (see Fig. 4). Follow-up paired *t*-tests showed that the effect was due to differences in the experimental condition. Specifically, mixed-race morphs were less attractive than within-race morphs when they were judged after racial classification, one-tailed $t(31) = 1.87$, $p < .05$, $\eta_p^2 = .10$. In the control condition mixed-race morphs were nonsignificantly more attractive ($p < .3$).

In sum, Study 3 formally replicates the interaction pattern from Studies 1 and 2. In both cases, the attractiveness of blended faces declined when (and only when) the experimental situation required that perceivers first classify the faces in terms of their "components" (the families in Studies 1 and 2; racial groups in Study 3). We propose that both effects are the consequence of contextual changes to stimulus fluency. When a blended face becomes relatively difficult to classify in relation to its components, the affect elicited by this disfluency generalizes to the stimulus face itself.

Study 4

A remaining question left open by the first three studies is the nature of the evaluative response underlying attractiveness judgments. Because those studies relied on self-reports, they leave open the possibility that changes in the categorization only influence the overtly expressed evaluative judgment about the blended faces, but do not change the actual affect associated with them. For example, attractiveness judgments could be a proxy for a "cold", cognitive assessment that the stimulus fits

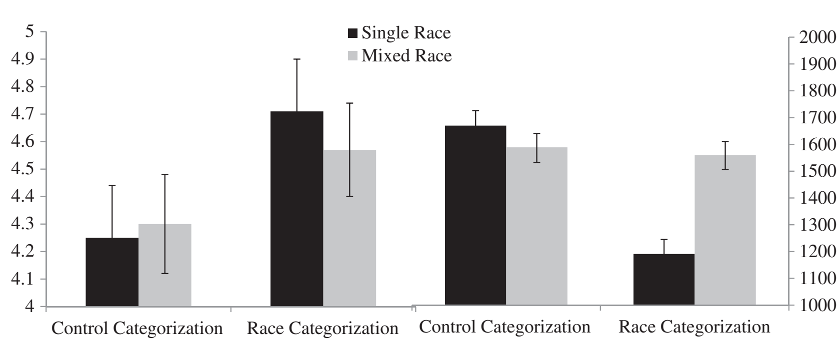

Fig. 4. Attractiveness ratings (left axis) and classification time (i.e., fluency; right axis) for single-race and mixed-race faces, as a function of experimental condition.

well or poorly into the provided category (e.g., is a "good" summary of seen faces, or is a "poor" example of a particular racial group).

To address this alternative interpretation, Study 4 gauged *nonverbal* affective responses to mixed-race faces using facial electromyography (EMG)—a well-validated indicator of affective response (Cacioppo, Petty, Losch, & Kim, 1986). Many studies have shown that positive reactions, even to mildly affective stimuli, manifest themselves in incipient smiles, as reflected by increased EMG activity over the cheek region, whereas negative affective responses manifest themselves in incipient frowns, as reflected by increased EMG activity over the brow region (Cacioppo, Bush, & Tassinary, 1992). Importantly, facial EMG has been successfully used in previous research on affective consequences of fluency, with processing ease enhancing zygomaticus activity and disfluency diminishing it (Harmon-Jones & Allen, 2001; Winkielman & Cacioppo, 2001, Winkielman et al., 2006).

Method

Participants were 12 undergraduates at the University of California San Diego who received course credit (4 additional participants did not provide usable EMG data). Stimuli used for the assessment of affective reactions via EMG measurement were 36 multi-face morphs. Single-race stimuli were 12 blends of 5 distinct Chinese faces or 12 blends of 5 distinct Caucasian faces. Mixed-race stimuli were 12 blends of these 5-face blends. These morphs appeared in blocks of 6 stimuli each. In addition, before each EMG assessment block, participants categorized 10 individual faces (5 Asian and 5 Caucasian, 2000 ms each), without making any evaluative judgments. The experimental group classified each face by race (by pressing "Z" and "M" keys). The control group pressed the "F" key as soon as a face appeared on the screen. In the following EMG assessment block the procedure was as follows. First, participants saw a 500 ms fixation cross (+). Then a face appeared for 2000 ms, and participants were asked to make a categorization response "quickly, without sacrificing accuracy." The experimental group again classified each face by race, and the control group pressed the "F" key as soon as a face appeared on the screen. After 2000 ms, the face was replaced by a fixation cross (+) for 3000 ms. Finally, participants made an attractiveness judgment. EMG was recorded during both the 500 ms pre-stimulus baseline period and the 5000 ms after stimulus onset over the zygomaticus (smiling) and corrugator (frowning) muscles. Stimulus presentation was controlled by E-Prime, and psychophysiological signals were recorded using a Biopac MP150 system. EMG data were processed in accordance with techniques used in previous research (Winkielman & Cacioppo, 2001; Winkielman et al., 2006).

Results and discussion

Fig. 5 shows baseline-corrected standardized activity over the zygomaticus "smiling" region. A planned contrast confirmed the hypothesized interaction. Specifically, participants smiled *less* to mixed-

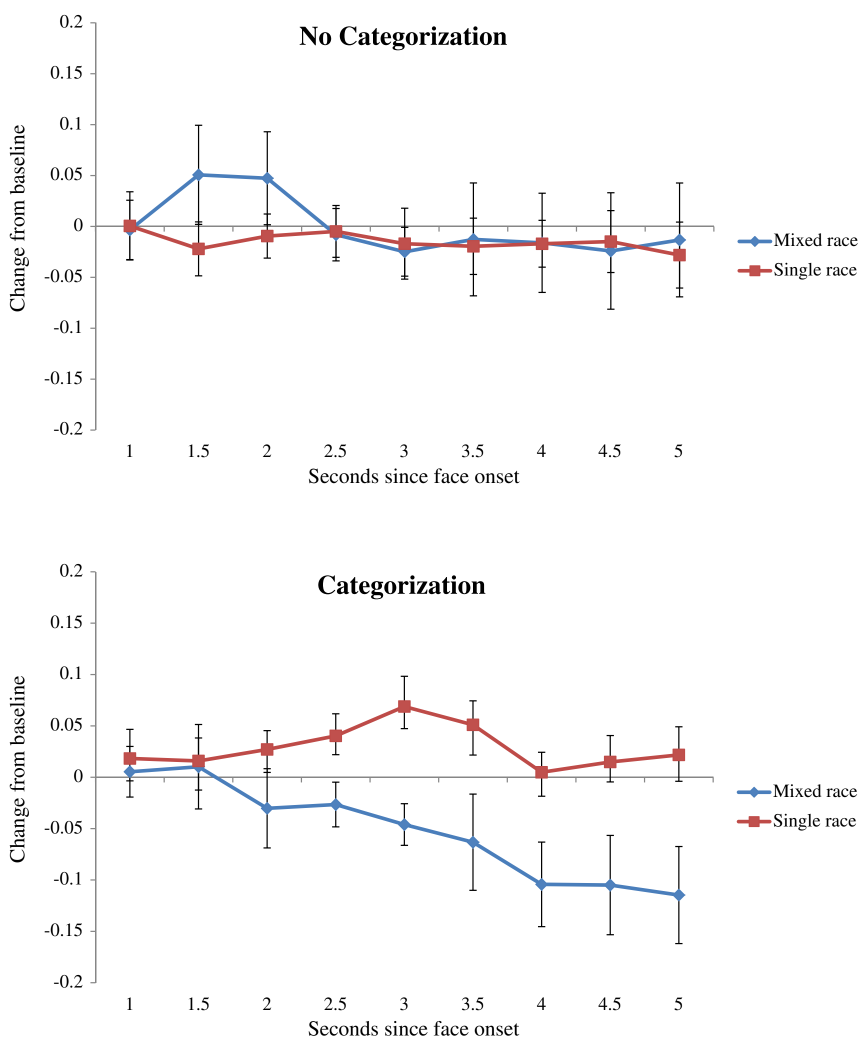

Fig. 5. Zygomaticus activity following exposure to single-race and cross-race faces, as a function of experimental condition. Values represent changes from baseline of standardized (z) scores.

race individuals who were classified by race (experimental condition), but *more* to mixed-race individuals when classification was not required (control condition), $t(10) = 2.51$, $p < .01$, $\eta_p^2 = .39$. Paired *t*-tests verified that the effect was due to differences in the experimental condition: Participants smiled less to mixed-race than to within-race individuals after classifying them by race (Ms = −.05 versus .02, SEs = .02 and .01), one tailed $t(6) = 2.66$, $p < .01$, $\eta_p^2 = .54$, but this was not the case in the control condition, $p < .2$. We found no effects in the corrugator region, associated with negative affect. This is consistent with previous work on fluency using EMG (Harmon-Jones & Allen, 2001; Winkielman & Cacioppo, 2001; Winkielman et al., 2006). However, the negative conclusions from the current study are limited by its small sample size (as are interpretations of the effect size estimates). Still, in sum, this study provides at least preliminary evidence against the interpretation that categorization-related decreases in attractiveness judgments observed in Studies 1–3 reflect only a "cold" assessment that the stimulus does not fit well into the provided category. Instead, the EMG data suggest that categorization can change genuine affective responses to mixed-race faces, reducing perceivers' positivity when such stimuli are difficult to process.

General discussion

Together, the current four studies represent the first experimental evidence that the attractiveness of faces, and in particular blended faces, depend on how — and how easily — they can be classified. Studies 1 and 2 make the point using laboratory-created morphs, whose attractiveness was significantly attenuated (albeit not entirely eliminated) when participants were forced to classify them in terms of their "parents" (the constituent categories). Studies 3 and 4 extend the results to "real" morphs — i.e., bi-racial faces — which were rated as less attractive, and elicited less positive physiological reactions, when they were first classified in terms of their constituent racial groups.

We propose that both sets of effects can be understood in terms of processing fluency. Faces, like any stimulus, are multiply categorizable and can differ in how *typical* they are of the various categories to which they belong. Facial morphs, in particular, are simultaneously good examples of the general category of "faces," but *also* poor examples of the parent faces from which they were generated. As such, morphs should be relatively fluent (and relatively attractive) in the context of the more general categorization task, and relatively disfluent (and relatively unattractive) in the context of the parent-related categorization task. The current studies establish the hypothesized role of processing fluency both statistically and experimentally. Specifically, participants in Studies 1 and 2 found family classification more difficult for strongly blended faces than weakly blended faces, and the extent of this disfluency predicted the morphs' relative attractiveness. Study 2 directly manipulated the role of processing fluency while holding constant the *context* of this categorization. That is, in Study 2, morphs were always seen as examples of their parents, because parent categories were explicitly provided. However, this experiment eliminated, in the control condition, the subjective effort needed to classify them. Only when participants themselves classified the faces did classification reduce the attractiveness of the blends.

From a theoretical perspective, these studies provide the first direct experimental evidence that facial attractiveness is partially fluency-driven. Moreover, the robust attractiveness of blended faces can be attenuated (and even reversed) by changing their implicit category membership. Therefore, these results provide a demonstration, and a plausible, mechanistic explanation, of the context-dependency of the "beauty-in-averageness" effect. The current studies also supplement the recent evidence that some morphs — i.e., blends whose "celebrity" parents are highly recognizable — are relatively *unattractive* (Halberstadt et al., 2013). Though Halberstadt et al. could only speculate about the cognitive mechanism underlying this reversal, the current results suggest that this "unattractiveness-in-averages" effect is due to disfluency elicited by competing person classifications (which occurs when the parents are highly recognizable in the morph).

More generally, the current studies are consistent with recent evidence that categorization can moderate even low-level facial processing, such as the configural integration underlying face inversion and cross-race effects (Ge, Wang, McCleery, & Lee, 2006; Hugenberg, Miller, & Claypool, 2007; Michel, Corneille, & Rossion, 2007). Together, this work illustrates that face perception in general, and the appeal of prototypicality in particular, cannot be solely understood via consideration of objective stimulus quality and low-level biological mechanisms designed to facilitate mate selection (Perrett, May, & Yoshikawa, 1994; Symons, 1979). Instead, it must be supplemented, at least in the case of faces, by consideration of contextual, high-level cognitive factors that take into account the organism's category structure and its flexible current tasks and goals.

The notion of context-dependent processing fluency also offers a cognitive resolution to the ambiguity surrounding how bi-racial faces (and individuals) are perceived. For instance, multi-dimensional scaling of computer-generated Caucasian and African–American faces has revealed that facial attractiveness increases with proximity to its own (same-race) prototype (Potter & Corneille, 2008). However, this finding conflicts with other evidence for the beauty of mixed-race blends (Lewis, 2010; Rhodes et al., 2005). The apparent discrepancy is resolvable, however, on the assumption that faces in the two paradigms are judged with respect to different implicit categories, such that bi-racial faces are only attractive when implicitly judged as "faces" (rather than "Caucasian faces" or "African–American" faces). From an applied perspective, this possibility suggests, counterintuitively, that when an individual's racial background in ambiguous, drawing attention to that ambiguity (or even to race itself) could engender less positive responses. That is, mixed-race individuals should elicit more positivity when race is *less* salient and, ironically, attention to racial background could, via the disfluency it engenders, *reduce* positive feelings towards them. Of course, these stronger claims about the role of fluency, and its relative importance compared to other factors, in racial attractiveness judgments remain speculative pending further evidence.

The current research provides the first empirical support for the categorization-driven changes to facial blends, along with a cognitive mechanism (classification difficulty) to account for them. Nevertheless, this account does not preclude the role of motivational factors. Indeed, a large Social Identity literature documents the importance of the salience of social groups in many cognitive processes, including the categorization and judgment of bi-racial faces (e.g., Brewer, 1988; Castano, Yzerbyt, Bourguignon, & Seron 2002; Eberhardt, Dasgupta, & Banaszynski, 2003; Shriver, Young, Hugenberg, Bernstein, & Lanter, 2008, etc.). However, these approaches (i) cannot be easily applied to the current Studies 1 and 2 (in which target race did not vary) and (ii) cannot explain recent demonstrations of the reversal of "beauty-in-averageness" for well-known celebrity morphs (Halberstadt et al., 2013). It is likely that the relation between the perceiver and target race would influence the appeal of racial blends, independent of (or perhaps interacting with) cognitive factors. More generally, the intersection of cognitive and motivational accounts of intergroup classification and judgment is a highly fertile area for future research.

References

Berntson, G. G., & Cacioppo, J. T. (2009). Evaluative processing. In D. Sander, & K. Scherer (Eds.), *Oxford companion to emotion and the affective sciences*. New York: Oxford University Press.

Brandell, J. R. (1988). Treatment of the bi-racial child: Theoretical and clinical issues. *Journal of Multi-cultural Counseling and Development*, *18*, 176–187.

Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull, & R. Wyer (Eds.), *Advances in social cognition*, *vol. 1*. (pp. 1–36)Hillsdale, NJ: Earlbaum.

Cacioppo, J. T., Bush, L. K., & Tassinary, L. G. (1992). Microexpressive facial actions as a function of affective stimuli: Replication and extension. *Personality and Social Psychology Bulletin*, *18*, 515–526.

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, *50*, 260–268.

Campbell, M. E., & Herman, M. R. (2010). Politics and policies: Attitudes toward multiracial Americans. *Ethnic and Racial Studies*, *33*, 1511–1536.

Castano, E., Yzerbyt, V. Y., Bourguignon, D., & Seron, E. (2002). Who may enter? The impact of in-group identification on in-group/out-group categorization. *Journal of Experimental Social Psychology*, *38*, 315–322.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*, 109–128.

Eberhardt, J. L., Dasgupta, N., & Banaszynski, T. L. (2003). Believing is seeing: The effects of racial labels and implicit beliefs on face perception. *Personality and Social Psychology Bulletin*, *29*, 360–370.

Etcoff, N. (2000). *Survival of the prettiest.* New York: Anchor Books.

Fiske, S. T. (1982). Schema-triggered affect: Applications to social perception. In M. S. Clark, & S. T. Fiske (Eds.), *Affect and cognition: The Seventeenth Annual Carnegie Symposium on Cognition* (pp. 66). Mahwah, NJ: Erlbaum.

Galton, F. (1879). Composite portraits, made by combining those of many different persons into a single resultant figure. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *8*, 132–144.

Ge, L., Wang, Z., McCleery, J. P., & Lee, K. (2006). Activation of face expertise and the inversion effect. *Psychological Science*, *17*, 12–16.

Gibbs, J. T., & Moskowitz-Sweet, G. (1991). Clinical and cultural issues in the treatment of biracial and bicultural adolescents. *Families in Society*, *72*, 579–591.

Halberstadt, J. B. (2006). The generality and ultimate origins of the attractiveness of prototypes. *Personality and Social Psychology Review*, *10*, 166–183.

Halberstadt, J. B., Pecher, D., Zeelenberg, R., Ip Wai, L., & Winkielman, P. (2013). Two faces of attractiveness: Making beauty-in-averageness appear and reverse. *Psychological Science*, http://dx.doi.org/10.1177/0956797613491969.

Harmon-Jones, E., & Allen, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual difference approaches. *Personality and Social Psychology Bulletin*, *27*, 889–898.

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Catergorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology*, *43*, 334–340.

Jackman, C. F., Wagner, G. W., & Johnson, J. T. (2001). The attitudes toward multiracial children scale. *Journal of Black Psychology*, *27*, 86–99.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, *1*, 115–121.

Lewis, M. B. (2010). Why are mixed-race people perceived as more attractive? *Perception*, *39*, 136–138.

Michel, C., Corneille, O., & Rossion, B. (2007). Race categorization modulates holistic face encoding. *Cognitive Science*, *31*, 911–924.

Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature*, *368*, 239–242.

Potter, T., & Corneille, O. (2008). Locating attractiveness in the face space: Faces are more attractive when closer to *their* group prototype. *Psychonomic Bulletin & Review*, *15*, 615–622.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*, 45–48.

Rhodes, G., Lee, K., Palermo, R., Weiss, M., Yoshikawa, S., Clissa, P., et al. (2005). Attractiveness of own-race, other-race, and mixed-race faces. *Perception*, *34*, 319–340.

Rhodes, G., Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only because of their symmetry? *Psychological Science*, *10*, 52–58.

Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science*, *7*, 105–110.

Rubenstein, A. J., Kalakanis, L., & Langlois, J. H. (1999). Infant preferences for attractive faces: A cognitive explanation. *Developmental Psychology*, *35*, 848–855.

Rubley, R. (2010). *Morpheus (Version 3.1) [Computer Software].* Santa Barbara, CA: Morpheus Software LLC.

Sanchez, D. T., & Bonam, C. M. (2009). To disclose or not to disclose biracial identity: The effect of biracial disclosure on perceiver evaluations and target responses. *Journal of Social Issues*, *65*, 129–149.

Saulny, S. (2011a, January 30). Black? White? Asian? More Young Americans Choose All of the Above. *The New York Times,* p.1.

Saulny, S. (2011b, March 24). Census data presents rise in multiracial population of youths. *The New York Times,* p.3.

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the face: Social context modulates the cross-race effect in face recognition. *Personality and Social Psychology Bulletin*, *34*, 260–274.

Symons, D. (1979). *The evolution of human sexuality.* New York: Oxford Univ. Press.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation increases positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.

Winkielman, P., Halberstadt, J. B., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science*, *17*, 799–806.

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch, & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Mahwah, NJ: Lawrence Erlbaum.

Winn, N. N., & Priest, R. (1993). Counseling biracial children: A forgotten component of multicultural counseling. *Family Therapy*, *20*, 29–35.

Zajonc, R. B. (1998). Emotions. (4th ed.)In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology*, *vol. 2*. (pp. 591–632)Boston, MA: McGraw-Hill.